Marine Biological Laboratory
WOODS HOLE, MASSACHUSETTS

July 6
~~June 30,~~ 1949

Dear Josh:

Thanks for your letter, which I take very seriously. I violently disagree with assertion that g is difficult to compute; a Pickett & Eckel model 4 slide rule will convert 1 datum/5 sec. At low S, g and S are the same, so you don't have to deal with anything off scale.

It is too bad that finished figs. were not ready to send. The superiority of the g plot is evident when shown graphically and compared with S plots of the same data. Your point that g should be defined as an experimental statistic I think is well taken and I plan to add something to emphasise this point. This leads me to an impasse in doing something about your suspicion that the g plot may not give a linear relation to dose when there is a Poisson distribution of units regardless of whether the S plot has approximated linearity, since the whole purpose of the derivation (13)-(14) is to show under what conditions the experimental statistic, g, is equal to np. Perhaps this could be shown as follows: Where p is the probability that a unit survives a dose D:

(1) $$p = e^{-kD}$$

Where there is an average of n units per organism np is the average number of surviving units per organism and,

(2) $$np = ne^{-kD}$$

If the units have a Poisson distribution and P_m is the proportion of organisms in which m units survive out of a total of n units;

(3) $$P_m = e^{-np}\frac{(np)^m}{m!}$$

and,

(4) $$P_0 = e^{-np}$$

Where S is the experimental survival of organisms, we assume according to the multi-unit hypothesis that:

(5) $$(1-S) = P_0$$

Now assume that in a given experiment we do not know whether equation (4) holds, i.e. we do not know the type of distribution of our hypothetical units at D=0 nor do we know the type of distribution of assumed surviving units at any dose. How much can we find out about this from the data, and what is the easiest way to do it?

One way is to just look at the data, or perhaps the S plot. I daresay that with three months of intensive training an intelligent individual could be taught to make fairly good guesses this way. You will admit that it is easier to temporarily assume that equation (4) holds, set 1-S equal to e^{-np}, compute np, then see where in the experiment lognp has a linear relation to dose. If we were able somehow to know the values of np in an experiment, it is evident from equation (2) above that log np would be found to have a linear relation to dose. Therefore if such a linear relation is found for np computed from (4) above, we are justified in thinking that our method of computing np was correct. Also, where this relation is non-linear we know that hthe theoretical np is an upper limit.

It may help to point out p is a free variable in (3) above, and is the only one which is a function of dose. Obviously the substitution of any of the possible values of p (i.e. the delivery of any dose) in (3) cannot change the form of the distribution, and hence the validity of (4), but merely shifts the expected (mean) value. This can be clearly seen in your derivation of the Delbruck equation where shifting q within its boundaries can have no effect on the derivation.

With constant initial n the error in deriving (14) from (13) depends on $1-p \approx e^{-p}$, and not on the value of n. For a given p, S is a function of n. Therefore the survival at which (14) holds is higher for higher n. It is evident on plotting "g" curves for various assumed distributions of initial n, that the approximation becomes good more rapidly for distributed n than for constant n, but I don't see the necessity for attempting a rigorous proof of this. In the light of your criticism I will try to further clarify the meaning of "g". Also, I would like your permiission to use your truncation correction and derivation of the Delbruck equation.

see table over

We wish you could come to Woods Hole, it's cool Here.

Kim

$1-p$	e^{-p}	% error
.3	.496	65.5
.4	.549	37.2
.5	.607	21.4
.6	.670	11.6
.7	.740	5.7
.8	.818	2.25
.9	.905	.056
.92	.923	.033

← This should be better than most experimental errors. ($p = .2$)

(approximate values)

n	$1-(1-p)^n$	e^{-np}
1	.20	.18
2	.36	.33
3	.49	.45
4	.59	.55
5	.67	.63
6	.73	.70
7	.79	.75
8	.83	.80
9	.87	.83
10	.89	.86